Improving the analytical assessment of fish stocks by providing parameters of data quality via InterCatch

Jose Castro, Eva Velasco, Paz Sampedro, Jose Rodríguez and Laurent Dubroca

INTRODUCTION

DCF (2008/949/EC) stated the provision of precision levels and sampling intensities of the estimates at national level. However, they are not regularly used in stock assessment.

Working Group on Data Needs
for Assessments and Advice
PGDATA 2016

INTRODUCTION

This presentation explores the potential of the tools currently available to facilitate the calculation of **data quality parameters** and its incorporation to the stock assessments:

1. Estimation.
2. Provision (InterCatch web-based system).
3. Use (a variety of assessment models).

1 - Estimation

The COST project (FISH/2006/15–Lot 2) developed an open source tool consisting on R packages for raising and estimating properties of statistical estimates derived from the DCF. The project established, in collaboration with the FishFrame existing format, a data exchange format for sampling, landings and effort from commercial fisheries (Jansen *et al., 2009)*.

This has served as the basis for the current Regional Data Base (RDB). The original COST library has been recently updated by the FishPi project.

- COSTeda package → to validate data taking into account its statistical quality (e.g. outlier detection).
- COSTdbe package → dealing with design-based estimates: sampling statistics, variance, confidence intervals, and CV .

2 - Provision

InterCatch:

1. HI matrix: No fields for data quality information.
2. SI matrix:
   - Field 21: “*varCATON*”.
   - Fields 22/23/24: “*InfoFleet*”; “*InfoStockCoordinator*”; “*InfoGeneral*”.
3. SD matrix:
   - Field 18: “*SampledCatch*”.
   - Fields 19/21: “*NumSamplesLngt*”; ”*NumSamplesAge*”.
   - Fields 20/22: “*NumLngtMeas*”; ”*NumAgeMeas*”.
   - Fields 31/32/33: “*varNumLanded*”; “*varWgtLanded*”; “*varLgtLanded*”.

3 - Use

Analytical assessment models used in Iberian stocks:

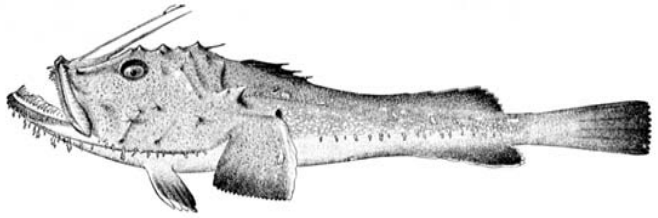
- **XSA** (megrim and 4-spot megrim): age-based model. Catch is assumed to be known exactly, and so the estimated observation error on these is zero.
- **AMISH** (horse mackerel): age-based model where the modeled population numbers-at-age are fit through an observation model for parameter estimation via a penalized likelihood. Uses sample size for commercial data.
- **GADGET** (southern hake): length-based model that compares modeled and ''real'' data to get a goodness-of-fit likelihood score, which can be weighed base on data quality information.
- **SS3** (length-based for white anglerfish, and age-based for sardine): allows use quality data parameters, CV's for abundance indices (surveys or LPUEs), and a sample size weighting input for commercial sampled data.

A crucial problem in modern stock assessments that combine multiple sources of data is how to weight each type of information.

Application: anp-8c9a assessment

Assessment of the white anglerfish southern stock (**apn-8c9a**):

- ICES stock data category: 1.
- Assessment type: Length-based model (SS3) that uses landings in the model and in the forecast.
- Input data: Landings, abundance indices (one survey index and two commercial indices) and length distribution .
- Discards: not quantified, assumed to be negligible.
- Working group: WG for the Bay of Biscay and the Iberian Waters Ecoregion (WGBIE).



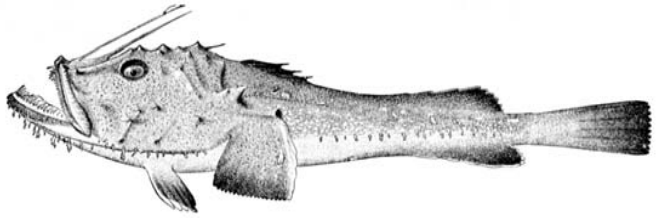
Application: anp-8c9a assessment

CV's of LPUE's for commercial data can be calculated by COST; however, LPUE are provided via Accessions, not via InterCatch.

- Input data: Landings, abundance indices (one survey index and two commercial indices) and length distribution.

SS3 Input sampling size → Effective sampling size

Methot & Wetzel, 2012



Application: calculation of m_{eff}

COST software does not provide m_{eff}, so it must be specifically calculated (Penington *et al*., 2002):

$$\mathrm{var}(\hat{R}) = \sum_{i=1}^{n} \frac{(M_i/\overline{M})^2 (\hat{\mu}_i - \hat{R})^2}{n(n-1)},$$

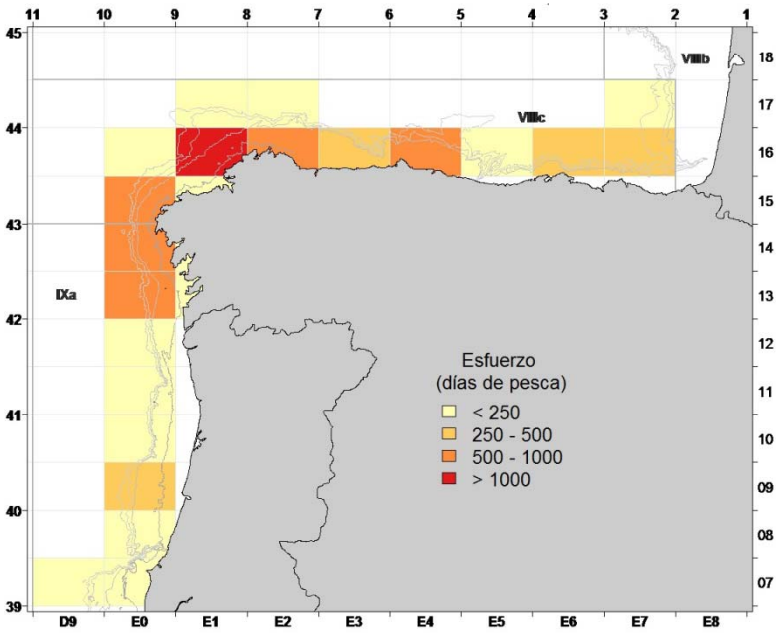
$$\hat{\sigma}_x^2 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m_i} (M_i/m_i)(x_{i,j} - \hat{R})^2}{M-1}$$

$$m_{eff} = \frac{\sigma_x^2}{var(R)}$$

m_{eff}: the number of fish that would need to be sampled at random so that the sample mean would have the same precision as an estimate based on a sample of n clusters (trips).

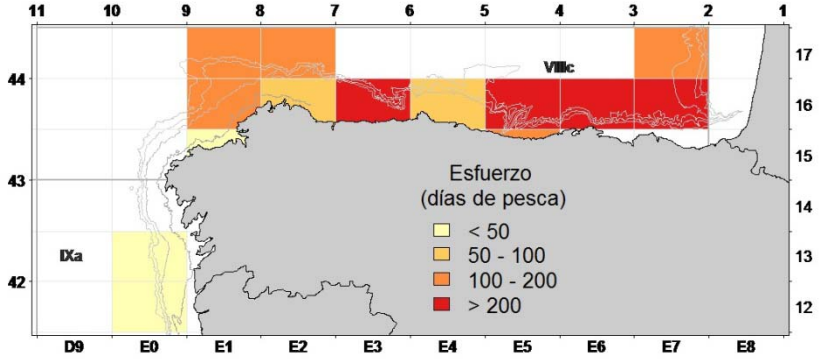
Application: métiers

2 Spanish métiers used in the anp-8c9a assessment:



OTB_DEF_>=55_0_0: Bottom otter trawl using cod end up 70mm to target demersal fish in ICES Divisions VIIIc-IXa.

GNS_DEF_>=100_0_0: Set gillnet using mesh size up 280mm to target anglerfish mainly in ICES Division VIIIc.



Application: input data

In last assessment (WGBIE 2016), m_{eff} were assumed to be 125 for both Spanish métiers, based on the expert knowledge.

Estimated values now:

Métier	Quarter	2011	2012	2013	2014	2015
OTB_DEF>=55	Q1	26.49	64.41	24.44	67.57	22.54
	Q2	30.34	88.23	29.20	64.36	43.69
	Q3	34.27	29.27	29.18	363.68	51.36
	Q4	84.16	38.62	39.86	48.45	38.43
GNS_DEF>=100	Q1	522.99	63.98	77.70	425.46	152.90
	Q2	230.26	132.83	292.77	227.46	532.84
	Q3	99.60	468.14	291.50	544.26	187.10
	Q4	156.94	343.81	297.66	555.56	281.90

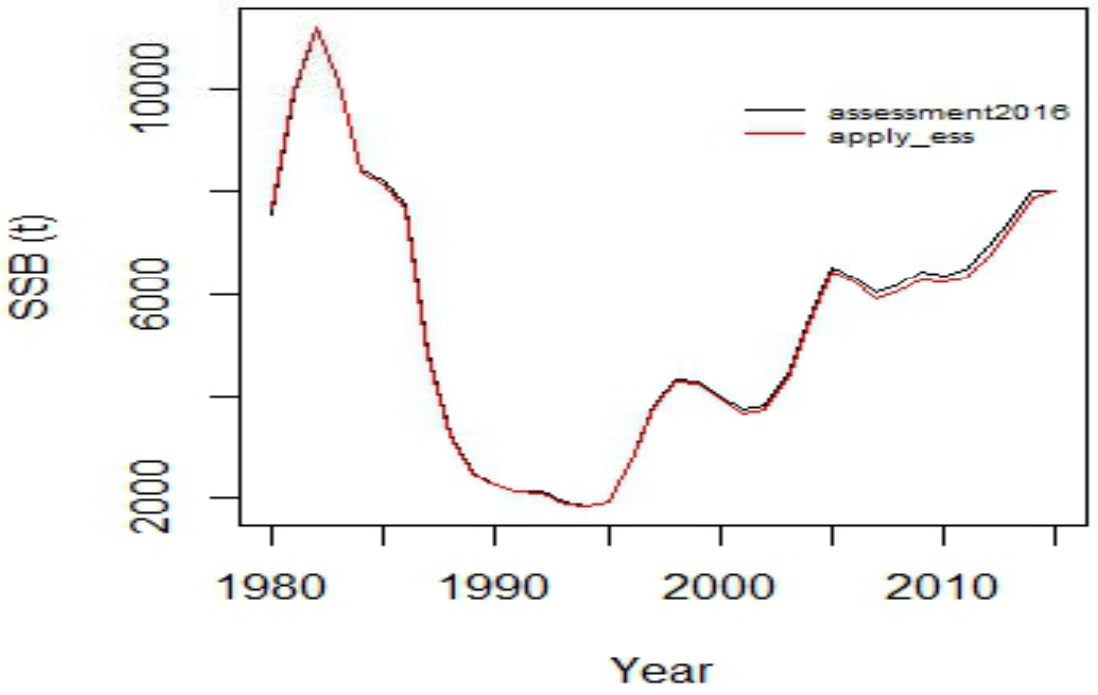
effective sample size (m_{eff}) is used as the input sample size for size composition.
Size composition records with effective sample sizes >150 were set to 150.

RESULTS: comparing assessments

	New run					WGBIE2016						
Year	Recruitment_ess	Biomass_ess	SSB_ess	F_ess	F/Fmsy_ess	Rec_base	Biomass_base	SSB_base	F_base	F/Fmsy_base	SSB_dif	F_dif
1980	424	13653	7698	0,33	1,06	425	13554	7566	0,32	1,03	-132	-0,01
1981	1683	15246	9984	0,33	1,06	1676	15216	9923	0,33	1,06	-61	0
1982	6726	14636	11192	0,38	1,23	6733	14658	11200	0,37	1,19	8	-0,01
1983	2921	13621	10153	0,51	1,65	2934	13671	10193	0,51	1,65	40	0
1984	801	13517	8393	0,54	1,74	797	13578	8445	0,54	1,74	52	0
1985	1708	12823	8146	0,56	1,81	1695	12903	8223	0,55	1,77	77	-0,01
1986	6009	10743	7678	0,83	2,68	5993	10829	7765	0,83	2,68	87	0
1987	4074	7390	4794	0,96	3,1	4061	7456	4863	0,96	3,1	69	0
1988	1630	7341	3242	1,49	4,81	1631	7384	3291	1,48	4,77	49	-0,01
1989	3018	5751	2456	1,23	3,97	3002	5772	2485	1,22	3,94	29	-0,01
1990	2392	4737	2239	0,9	2,9	2399	4752	2257	0,89	2,87	18	-0,01
1991	921	4652	2105	0,88	2,84	921	4661	2117	0,88	2,84	12	0
1992	1171	4404	2092	0,93	3	1169	4414	2103	0,92	2,97	11	-0,01
1993	1396	3516	1891	0,7	2,26	1391	3526	1902	0,69	2,23	11	-0,01
1994	2891	3343	1840	0,6	1,94	2890	3354	1851	0,6	1,94	11	0
1995	2154	3882	1920	0,4	1,29	2165	3895	1932	0,39	1,26	12	-0,01
1996	448	5741	2739	0,43	1,39	451	5763	2756	0,43	1,39	17	0
1997	208	6840	3794	0,48	1,55	208	6877	3823	0,48	1,55	29	0
1998	181	6276	4287	0,4	1,29	181	6327	4330	0,39	1,26	43	-0,01
1999	481	5345	4235	0,3	0,97	481	5402	4288	0,3	0,97	53	0
2000	572	4674	3952	0,25	0,81	570	4734	4010	0,25	0,81	58	0
2001	3163	4411	3663	0,19	0,61	3164	4470	3722	0,19	0,61	59	0
2002	1576	5127	3747	0,2	0,65	1590	5191	3807	0,2	0,65	60	0
2003	395	7199	4327	0,31	1	397	7269	4392	0,31	1	65	0
2004	1751	8609	5454	0,34	1,1	1747	8696	5533	0,33	1,06	79	-0,01
2005	1116	8908	6438	0,39	1,26	1129	9009	6531	0,38	1,23	93	-0,01
2006	1357	8427	6227	0,37	1,19	1364	8532	6327	0,37	1,19	100	0
2007	583	8057	5899	0,31	1	587	8173	6006	0,31	1	107	0
2008	493	8191	6052	0,29	0,94	516	8319	6169	0,29	0,94	117	0
2009	710	8091	6290	0,29	0,94	725	8234	6419	0,29	0,94	129	0
2010	1074	7644	6235	0,21	0,68	1034	7807	6376	0,21	0,68	141	0
2011	1079	7766	6328	0,17	0,55	1038	7948	6488	0,16	0,52	160	-0,01
2012	531	8506	6736	0,18	0,58	457	8680	6919	0,18	0,58	183	0
2013	638	9230	7250	0,18	0,58	640	9355	7428	0,18	0,58	178	0
2014	1321	9745	7899	0,23	0,74	1181	9772	8015	0,23	0,74	116	0
2015	190	9701	8011	0,2	0,65	178	9596	8008	0,21	0,68	-3	0,01
2016			8094					7941			-153	

RESULTS: comparing assessments

Comparison between last WGBIE assessment and the new run using m_{eff}:



RESULTS: comparing assessments

Comparison between last WGBIE assessment and the new run using m_{eff}:

Parameters	WGBIE	New run
F30–130cm (2016)	0.21	0.20
SSB (2017)	7.984 kt	8.252 kt
Catch (2017)	2253 t	2.431 t

CONCLUSION: regular provision

Specific

The comparison of assessment results indicate that weighting process might have an effect in the estimate of stocks status indicators SSB, F and Recruitment.

Meff - calculated and provided by metier and quarter - improved the previos sampling size input which was based on expertise knowledge criteria and was somehow arbitrary.

General framework

Using Intercatch to provide quality estimates ensures that this information is available at the same time as the data themselves.

Direct communication and agreement with the end-user to provide quality data according to the specific assessment model would improve the results.